Appendix B
Caustic Surfaces

A caustic surface or "burning curve" in geometric optics is a boundary separating accessible and inaccessible regions for a given family of light rays. The rays within that defined family can "pile up" against the boundary but they never cross it. The boundary, therefore, is an envelope with respect to that family. Caustics are very common in everyday life—for example, the double crescent-shaped cusp of reflected light from a point source at the bottom of a coffee cup or inside a wedding band.

Another caustic phenomenon is the rainbow [1], where for a given color the scattering angle from the raindrop assumes a stationary value with respect to the impact parameter or distance that the incident ray makes relative to the center of the raindrop (at a local minimum of about 138 deg for the primary and a local maximum of about 130 deg for the secondary). Since light rays are equally likely to impinge on a raindrop at any impact distance, stationarity of the scattering angle (which occurs for the primary at an impact distance of about 80 percent of the radius of the raindrop) means that the light rays "pile up" at that exit angle. And because their scattering angle and, therefore, their phase at exit from the raindrop are stationary with respect to the impact parameter at this angle, they all are essentially in phase and reinforce each other in the vicinity of this stationary point upon arrival at the observer. This results in the "caustic" phenomenon. Light rays from a raindrop at scattering angles slightly less than the local maximum of the primary rainbow can have impact distances that are slightly higher and lower than the impact distance that provides the stationary scattering angle. Because these higher and lower rays travel through the raindrop along slightly different paths, their travel times to the observer differ, which can result in their arriving at the observer both in and out of phase; both constructive and destructive interference can result. These are the supernumerary bands observed with some rainbows. Supernumerary bands in rainbows are the analog to the amplitude and phase variability that we observe in a radio occultation signal near a caustic.

The stationarity property of the contact point of a ray with its caustic surface defines the mathematical form of the caustic surface. A caustic surface is an envelope generated by a family of curves, which in the case of geometric optics are themselves stationary phase paths. For convenience, we use here a thin-screen framework to develop a mathematical description of a caustic surface. Let $y = f(x,h)$ describe a one-parameter family of ray paths, where y is the altitude of a point on the ray path, x is its perpendicular distance from the thin screen, and h is the parameter whose value identifies the family member. The boundary conditions are $y = h = f(0,h)$ at the thin screen and $y = h_{\mathrm{LG}} = f(D,h)$ at the low Earth orbiting (LEO) satellite. It is straightforward to work out the explicit functional form of $f(x,h)$ from the thin-screen relationship in Eq. (2.2-5). Thus,

$$y = f(x,h) = h - x\alpha\big(a(h)\big) \tag{B-1}$$

where in this model $a(h)$ provides the relationship between impact parameter and thin-screen altitude used to generate the bending angle α.

As we vary h, we generate the family of rays that satisfies the boundary conditions at $x = 0$ and at $x = D$. An envelope, if one exists, is defined by the condition that, at its contact point with a ray, it must be tangent to that member of the family. Also, there must be a continuum of contact points over at least a subset of the family of rays. Let the functional form defining the envelope be given by $y = g(x)$, and let the contact point be designated by $\left(x^{\dagger}, y^{\dagger}\right)$, which is an implicit function of the parameter h. At a contact point, we require that

$$y^{\dagger} = g\left(x^{\dagger}\right) = f(x,h)\big|_{x=x^{\dagger}} \tag{B-2}$$

As the contact point varies due to varying h, the first-order variations of g and f with respect to h are given by

$$\frac{dy^{\dagger}}{dh} = \frac{dg}{dx^{\dagger}}\frac{dx^{\dagger}}{dh} = \left(\frac{\partial f}{\partial x}\frac{dx}{dh} + \frac{\partial f}{\partial h}\right)_{x=x^{\dagger}} \tag{B-3}$$

The tangency condition on the ray and the envelope at the contact point requires that $dg / dx^{\dagger} = (\partial f / \partial x)_{x=x^{\dagger}}$. It follows that the position y of the ray member must be stationary with respect to the parameter h at the contact point. That is,

$$\left.\frac{\partial f}{\partial h}\right|_{x=x^{\dagger}} = 0 \tag{B-4}$$

These contact and stationarity conditions in Eqs. (B-2) and (B-4) enable one to solve for the functional form of the envelope, $y = g(x)$. The stationarity condition in Eq. (B-4) implies that $\partial y/\partial h = 0$ in Eq. (B-1), which is equivalent to letting the defocusing factor $\zeta \to \infty$, which in turn is equivalent to setting both the first and second partial derivatives of the Fresnel phase function [Eq. (2.5-1)] with respect to h to zero. At the contact point with the envelope, one obtains from Eq. (B-1)

$$x^{\dagger} = \left(\frac{d\alpha}{dh}\right)^{-1} \tag{B-5}$$

From this relation, one obtains $h^{\dagger} = h(x^{\dagger})$ through the definition of $\alpha(a(h))$, for example, by Eqs. (A-39) and (A-40) for Case C. It follows that

$$y^{\dagger} = f(x^{\dagger}, h^{\dagger}) = g(x^{\dagger}) = h(x^{\dagger}) - x^{\dagger}\alpha\left(a\left(h\left(x^{\dagger}\right)\right)\right) \tag{B-6}$$

which is the functional form for the envelope.

The caustic for Case C is shown (not to scale) as the limiting concave arc in Fig. B-1, which shows the local families of the ordinary (b) rays and the anomalous (a) rays. These are the "–" rays that are applicable below h_o. In this figure, h_o is the altitude where the discontinuity in the lapse rate occurs [see also Fig. 2-7(c)]. The LEO plane is on the left side of Fig. B-1; the thin screen is on the right side. The (a) family, whose rays begin from the thin screen in the altitude range $h_o \geq h_a \geq h(2)$, generates the envelope. These are the so-called anomalous rays. The ordinary family of "–" rays, the (b) rays, which originates

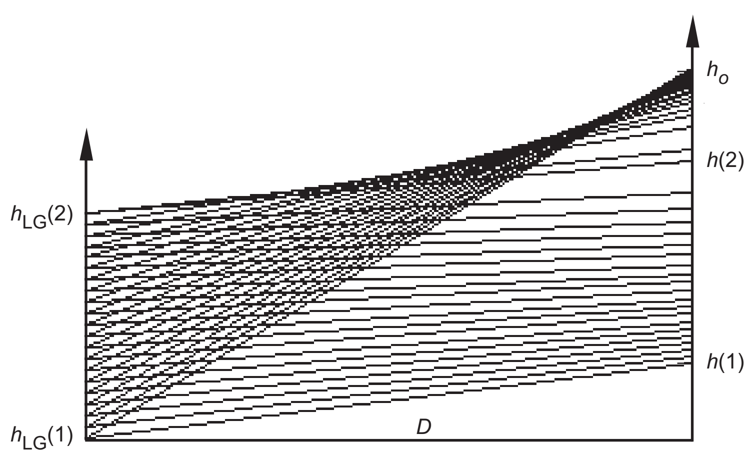

Fig. B-1. Limiting arc is the caustic surface generated by a family of anomalous rays.

from the thin screen in the altitude range $h(2) \geq h_b \geq h(1)$, does not generate an envelope. Figure B-1 shows that one ray from each of these families intercepts the h_{LG} plane at the same point when $h_{\mathrm{LG}}(1) \leq h_{\mathrm{LG}} \leq h_{\mathrm{LG}}(2)$. To improve the clarity in the figure, the local family of "+" rays originating from thin-screen altitudes above h_o is not shown. These are the (m) rays, the original ray system before impacting the discontinuity. Adding these (m) rays to the figure would produce a triplet system of rays arriving at the LEO when the altitude of the Global Positioning System (GPS)–LEO line is in the range $h_{\mathrm{LG}}(1) \leq h_{\mathrm{LG}} \leq h_{\mathrm{LG}}(2)$.

The first caustic contact is at $h^{\dagger} = h(2)$ with the GPS–LEO line at $h_{\mathrm{LG}} = h_{\mathrm{LG}}(2)$. This marks the birth of the ray systems (a) and (b). As the altitude of the GPS–LEO line migrates downward in the range $h_{\mathrm{LG}}(1) \leq h_{\mathrm{LG}} \leq h_{\mathrm{LG}}(2)$, the altitude of the (a) ray in the thin screen migrates upward in the range $h_o \geq h_a \geq h(2)$. The defocusing factor for the (a) ray is negative; hence, the appellation "anomalous." Concurrently, the altitude of the (b) ray migrates downward in the range $h(2) \geq h_b \geq h(1)$; its defocusing factor is positive. When the altitude of the GPS–LEO line drops below $h_{\mathrm{LG}}(1)$, only the (b) ray survives.

If a ray contacts the caustic surface along a tangential between the thin screen and the LEO, which every member of the anomalous branch does, then the Calculus of Variations tells us that this interior point of contact is a conjugate point. A ray with a conjugate point, although a path of stationary phase, is not necessarily a path along which the path delay is a local minimum; it can be, and in this case it is a local maximum, as Fig. 2-9(c) shows.[1] The anomalous branch violates the Jacobi condition from the Calculus of Variations, which requires that the ray have no interior contact point with its envelope. The Jacobi condition is an additional necessary condition that the ray path must satisfy to yield a local minimum in elapsed signal time, or phase delay. Caustics are well known in seismology [2].

---

[1] A simple example of a conjugate point is taken from the system of geodesics on a spherical surface. Here the stationary phase path is a great circle. Consider the family of great circle paths, all originating from the north pole and generated by varying the longitude parameter, for example, the longitude of the path when it makes its southward crossing of the equator. The distance from the north pole to any geographical location on the sphere is a global minimum along a great circle route provided that the path does not first pass through the south pole. If it does, there is an alternate family member, the great circle path with its longitude parameter 180 deg different, that provides a shorter distance to the same geographical location. The south pole is a conjugate surface for this family of geodesics originating from the north pole. Unfortunately, the conjugate surface in this example is somewhat pathological, shrinking to a single point at the South Pole.

References

[1] H. M. Nussenzveig, "The Theory of the Rainbow," *Scientific American*, vol. 236, no. 4, pp. 116–127, 1977.

[2] K. Bullen and B. Bolt, *An Introduction to the Theory of Seismology*, Cambridge, United Kingdom: Cambridge University Press, 1993.